ILLINOIS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN

PRODUCTION NOTE

University of Illinois at
Urbana-Champaign Library
Large-scale Digitization Project, 2007.

Technical Report No. 401

WHAT CAUSES CHILDREN'S FAILURES TO DETECT INCONSISTENCIES IN TEXT? REPRESENTATION VS. COMPARISON DIFFICULTIES

Stella Vosniadou, P. David Pearson, & Theresa Rogers
University of Illinois at Urbana-Champaign

January 1987

Center for the Study of Reading

TECHNICAL REPORTS

UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN
174 Children's Research Center
51 Gerty Drive
Champaign, Illinois 61820

CENTER FOR THE STUDY OF READING

Technical Report No. 401

WHAT CAUSES CHILDREN'S FAILURES TO DETECT INCONSISTENCIES IN TEXT? REPRESENTATION VS. COMPARISON DIFFICULTIES

Stella Vosniadou, P. David Pearson, & Theresa Rogers
University of Illinois at Urbana-Champaign

January 1987

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

This research was supported in part by the National Institute of Education under Contract No. NIE 400-81-0030. We wish to thank the children, teachers and principals of the Carey Busey School, Dr. Howard School, and Thomas Paine School in Champaign-Urbana, Illinois. We also wish to thank Anne Goudvis and Barbara Hunter for help in testing the children, and Rick Ommanson, Bob Reeve and Bill Brewer for helpful comments on earlier drafts. Portions of this paper were presented at the annual meeting of the American Educational Research Association, New Orleans, April 1984, and at the biennial meeting of the Society for Research in Child Development, Toronto, Canada, April 1985. Requests for reprints should be sent to Stella Vosniadou, Center for the Study of Reading, University of Illinois, 51 Gerty Drive, Champaign, Illinois 61820. This paper will appear in the Journal of Educational Psychology.

Abstract

This paper reports two experiments that investigated whether elementary school children's difficulties in detecting inconsistencies in text are related to their failure to represent each of two inconsistent propositions in memory or to their failure to compare the representations of the inconsistent propositions to each other once each has been represented in memory. Overall, the experiments considered three kinds of inconsistencies--*falsehoods* (a textual proposition conflicts with a potentially known fact), *factual contradictions* (one textual proposition conflicts with another textual proposition and one of these propositions is a potentially known fact) and *textual contradictions* (one textual proposition conflicts within a second textual proposition *and* neither is a known fact). In Experiment 1, first, third, and fifth grade children (N = 80) were asked to detect familiar falsehoods and unfamiliar factual contradictions in narratives. Results showed that the familiar falsehoods were easier to detect than the unfamiliar factual contradictions. In Experiment 2 (N = 30), however, when the familiarity variable was controlled, no differences in inconsistency detection were observed between falsehoods, factual contradictions, and textual contradictions. In addition, in both experiments an analysis of the recall protocols indicated that detection failures were related more to incomplete recall of the inconsistent information than to difficulty in comparing the inconsistent propositions. The results indicated that children's inconsistency detection failures are related more to difficulties in forming accurate mental representations of textual propositions than to difficulties in comparing the inconsistent information once it is represented in memory. It was suggested that greater attention should be paid to the conditions that facilitate text representation, since those conditions are likely to affect comprehension monitoring as well.

WHAT CAUSES CHILDREN'S FAILURES TO DETECT INCONSISTENCIES IN TEXT?

REPRESENTATION VERSUS COMPARISON DIFFICULTIES

During the last few years we have seen a rapid growth of research on children's awareness of inconsistencies or omissions in text under the general rubric of studies on comprehension monitoring. A number of studies have shown that elementary school children often indicate that they have understood highly ambiguous or contradictory materials (e.g., Flavell, Speer, Green, & August, 1981; Ironsmith & Whitehurst, 1978; Markman, 1977, 1979; Kotsonis & Patterson, 1980). Usually, young children are less sensitive to such errors than older children, and poor readers are less sensitive than good readers (August, Flavell, & Clift, 1984), although the problem is not confined to children. Junior high school students, particularly those identified as poor readers (Garner, 1980; in press), and even college students (Baker, 1979; Baker & Anderson, 1982) have problems evaluating texts for clarity and consistency.

In this paper we shall examine elementary school students' ability to detect inconsistencies in text. We shall examine three kinds of inconsistencies--falsehoods, factual contradictions, and textual contradictions.

A textual contradiction is detected when a student recognizes the inconsistency between two statements, such as those in examples (1) and (2), both of which occur in the text.

(1) Henry walked through the open door into the kitchen.

(2) Henry could not open the kitchen door because his key was bent.

The critical features of a textual contradiction are that (a) either proposition could be true or false, but (b) both cannot be true in the same context, and (c) neither can be subjected to a test of veracity independent of the text in which they occur. Obviously, they tend to occur in "stories": texts in which things that are true in the world of the text may not be true in the real world.

A factual contradiction is detected when a student recognizes the inconsistency between two statements, one of which is true while the other is false, when both occur in the text. Examples (3) and (4) comprise such a contradiction.

(3) Fish that live at the bottom of the ocean cannot see anything.

(4) Fish that live at the bottom of the ocean can see the color of their food (examples from Markman, 1979).

Factual contradictions are identical to textual contradictions in surface form only. Like textual contradictions, both statements cannot be true in the same context. Unlike textual contradictions, only one of the propositions can be true, because only one is true, and both can be subjected to tests of veracity independent of the text in which they occur. So a student who knew a lot about aquatic vision *could* use that knowledge to detect the inconsistency between (3) and (4). But even a student who knew next to nothing about aquatic vision could detect the inconsistency, in the same way that he or she would have to detect the inconsistency between the two propositions in a textual contradiction.

The contradiction in a falsehood is detected when a student recognizes the inconsistency between a proposition (the false statement--see example (5)) which comes from the text and some information (the truth--see example (6)) which comes from existing knowledge.

(5) Most people prefer not to eat their favorite foods (example from Markman & Gorin, 1981).

(6) What makes your favorite foods favorite is that you like to eat them more than you like to eat other foods.

What makes a falsehood different from a factual contradiction is that only one proposition (the falsehood) is present in the text. Notice also that from the point of view of a subject asked to look for inconsistencies, it makes sense to ask for falsehood detection only for relatively familiar and common "truths." For example, if young children were presented with a statement such as (7), few would ever even consider the possibility that it might not match something they already know.

(7) The average bank vault is larger than the average professorial office.

We have conceptualized the inconsistency detection task as involving the following subcomponents for either factual or textual contradictions:

(1) Read or listen, encode, and represent the propositions in working memory.

(2) Compare the representations of the inconsistent propositions to one another.

(3) Detect the inconsistency.

(4) Report it.

In the case of falsehoods, only one proposition needs to be subjected to step (1); however, falsehoods require an additional procedure in step (1) so that it would become,

(1a) Read or listen, encode, and represent the false proposition in working memory.

(1b) Retrieve the information that contradicts it from long-term memory.

Children's inconsistency detection failures have been attributed to difficulties with each subcomponent of this assumed inconsistency detection process. For example, some researchers have argued that children have problems encoding and representing the inconsistent information in memory (Stein & Trabasso, 1981; Wimmer, 1979), while others have focused on difficulties related to the comparison subcomponent of the inconsistency detection process (Markman, 1979, 1981, 1985). It has also been pointed out that children may fail to detect encoded inconsistencies because they do not know the criteria that define an inconsistency (Whitehurst, 1981; Flavell, 1981), or because they have difficulty reporting detected inconsistencies verbally (e.g., Baker, 1979; Flavell, Speer, Green & August, 1981; Patterson, Cosgrove, & O'Brien, 1980). In this study, we shall concentrate on the first two sources of inconsistency detection failures: representation and comparison.

With respect to comparison difficulties, Markman (1979, 1981) has argued that children judge their comprehension of text by evaluating the truth of individual propositions rather than by judging how internally consistent these propositions are with respect to each other, although she acknowledges that evaluation criteria may change as a result of training (Markman & Gorin, 1981; see also Elliot-Faust, & Pressley, 1986). According to Markman, children find it difficult to engage in the inferential and constructive processing which is often necessary to evaluate a text for internal consistency; they do not connect sentences, draw appropriate inferences, and--most important--they do not compare inconsistent propositions to each other.

The proposal that children have difficulty with the proposition comparison subcomponent of the inconsistency detection task was based on two findings in Markman's (1977, 1979) experiments: (1)

children were more likely to question the truth of individual propositions than the internal consistency of the texts, and (2) many of the children who did not detect the inconsistencies had nevertheless recalled all the relevant information necessary for their detection.

One limitation of Markman's experiments is that the materials employed contained information unfamiliar to young children. Since many cognitive skills are first exhibited in familiar contexts (Gelman, 1978), it could be argued that children would be more likely to detect inconsistencies in texts describing familiar information. In fact, there are a few studies which have shown that young children are capable of inconsistency detection when the content of the text is familiar and the task is appropriate (e.g., Ackerman, 1981; Raphael & Tierney, 1980; Stein & Trabasso, 1981; Wimmer, 1979).

Based on such findings, Wimmer (1979) and Stein and Trabasso (1981) have argued that children's inconsistency detection difficulties lie in the representational subcomponent of the inconsistency detection task, rather than in the comparison subcomponent. According to their proposal, children have difficulties remembering the inconsistent information, particularly when it is new and unfamiliar. When the inconsistent information is remembered, inconsistency detection is high (see Wimmer, 1979).

Familiarity with the subject matter described in a text can affect inconsistency detection in either the representation, or the comparison, subcomponents of the assumed inconsistency detection process, or both. We know from a host of studies that prior knowledge can influence subjects' ability to represent a linguistic input in memory (e.g., Brown, Smiley, Day, Townsend, & Lawton, 1977; Bransford & Johnson, 1973). Thus, familiar inconsistencies should be easier to represent than unfamiliar ones. Prior knowledge can also affect the process of comparing the inconsistent propositions. When a proposition is both inconsistent with a second proposition *and* violates a well-known fact, children may need only to compare their representation of the false proposition against existing knowledge, i.e., evaluate the proposition for truth, rather than consistency. This is not possible when the text contains two propositions which are both inconsistent and unfamiliar. In this latter case the two inconsistent propositions must be represented in memory and then must be compared to each other. The latter situation is likely to be more difficult than the former.

The primary purpose of the two experiments described in this paper was to further investigate these two sources of children's inconsistency detection failures--representation versus comparison. In order to do so, we compared children's detection of falsehoods versus textual and factual contradictions. The purpose of the first experiment was to evaluate the hypothesis that *familiar falsehoods* are indeed easier to detect than *unfamiliar contradictions*. First, third, and fifth grade children were asked to detect *familiar falsehoods* (violations of well-known facts) and *unfamiliar factual contradictions* (two relatively unfamiliar propositions which contradicted each other). We hypothesized that inconsistencies in the falsehoods would be more likely to be detected than those in the factual contradictions because they were easier both to represent and to compare. Recall that for a falsehood only *one* textual proposition, one that violates a well-known fact, must be represented in memory, whereas *two* textual propositions, describing relatively unfamiliar facts, must be represented in the case of a factual contradiction. Hence, familiar falsehoods must be easier to represent. Furthermore, other things being equal, familiar falsehoods require comparing the representation of one proposition against existing knowledge (for truth), whereas unfamiliar factual contradictions require the comparison of the representations of two textual propositions against each other (for internal consistency). According to Markman (1979), evaluating a proposition for truth is easier than comparing two propositions for internal consistency.

The second experiment was conceptualized following the completion of the first experiment in order to provide a more precise test of the source of inconsistency detection by controlling the familiarity variable that was allowed to vary in Experiment 1.

In both experiments, inconsistency detection was measured on the basis of the children's verbal responses. Children were asked to indicate what it was about the text that did not make sense and why. While children demonstrate better awareness of inconsistencies when non-verbal measures of inconsistency detection are employed (e.g., Flavell, et al., 1981), it is not altogether clear exactly what they are aware of. In addition, a number of factors other than verbal report difficulty may stand in the way between noticing and reporting an inconsistency. For instance, young children may fail to report detected inconsistencies because they are reluctant to attribute fault to an adult (e.g., Robinson & Robinson, 1976a, 1976b, 1977), or because they use prior knowledge to "repair" a detected inconsistency (e.g., Baker, 1979; Baker & Anderson, 1982; Winograd & Johnston, 1980). In order to minimize such "metacognitive" sources of error, the children were told explicitly that "there was something wrong with the way the author wrote the stories." Thus, the children did not have to decide whether the text contained false or inconsistent information themselves--they were told so; neither were they put in the position to attribute blame to its presumably adult author. Their task was to simply discover what it was about the stories that did not make sense.

EXPERIMENT 1

Experiment 1 investigated elementary school children's detections of falsehoods and factual contradictions. The falsehoods involved the *implicit* violation of an *everyday* scientific principle (i.e., the knowledge that sugar cannot lift itself up from the top of a table and go back in its box). The factual contradictions involved the violation of a *relatively* technical scientific principle (i.e., that magnets do not pick up sugar) when the principle was *explicitly* stated in the text. Thus, both the falsehoods and the contradictions negated a true fact about the world, but differed (a) in their explicitness and (b) in their familiarity.

The falsehoods negated knowledge firmly grounded in children's everyday experience with the world. A pilot study, which will be described later in greater detail, showed that the familiar falsehoods could be judged as false by 92 to 97% of a comparable sample of children. The factual contradictions, which negated relatively unfamiliar scientific principles, were known to be false by only 53 to 76% of a comparable sample of children.

The falsehoods were not explicitly contradicted in the text. It would be odd, if not silly, to assert that "sugar does not lift itself up in the air." Furthermore, the question of interest was whether the children would spontaneously evaluate the truth of these propositions by comparing them against existing knowledge. The factual contradictions were explicitly negated in the text. Their detection required comparing two inconsistent propositions for consistency, at least for those children who lacked knowledge of the principle.

It was hypothesized that a better detection rate would be obtained for falsehoods than contradictions, despite the explicit nature of the latter, because the falsehoods were easier to represent (being more familiar) than the contradictions, and did not require comparing two propositions for internal consistency.

Information regarding children's abilities to represent and compare the inconsistent propositions was obtained from an analysis of the recall protocols. If children had problems encoding and representing the unfamiliar information, we expected to find poor recall of the inconsistent propositions. If children had difficulty comparing the inconsistent information, we expected to find a low correlation between recall and inconsistency detection (like that obtained by Markman, 1979). In other words, many of the inconsistent propositions would be recalled, but the inconsistencies would not be detected.

A second issue pursued in this experiment was the relative difficulty of inconsistency detection in a listening vs. a reading comprehension task. Throughout the elementary school years, children's comprehension of text is better in listening than reading comprehension tasks (Sticht, Beek, Hanke, Kleiman, & Mamee, 1974; Pearson & Fielding, 1983). This finding has been explained on the grounds that reading involves an additional resource allocation problem for the young child not present in the listening situation; that is, more cognitive resources must be allocated to the task of decoding written text than decoding text presented orally, leaving fewer resources available to process the text for meaning. If inconsistency detection is affected by the processing requirements of the text, then we should expect the frequency of inconsistency detection to be lower in a reading than a listening task for elementary school children. This should be particularly true in the case of the contradictions when compared to falsehoods, since contradiction detection places even greater demands on the constructive processing of the information presented in the text.

Method

Subjects

Eighty elementary school students participated in the study. Twenty first-graders (ranging in age from 6.0 to 7.0, with a mean age of 6.5), 40 (20 in a reading condition and 20 in a listening condition) third-graders (ranging in age from 7.1 to 9.9, with a mean age of 8.7), and 20 fifth-graders (ranging in age from 10.2 to 11.3, with a mean age of 10.7). The children were drawn from two elementary schools in a mid-size Midwestern city. In each group approximately half of the children were girls and half were boys. All children read at or above their grade level as determined by scores on the reading section of the Iowa Test of Basic Skills, and/or by teacher judgment. Only competent readers were chosen in order to avoid a possible confounding of serious decoding problems within the reading condition.

Materials

There were four stories with the same basic structure; each contained an initiating event, the creation of a problem situation, and a set of actions leading, eventually, to its resolution. All the stories centered around the same basic theme: How the protagonist (a young girl) gets into a problem situation which she does or does not resolve. All the inconsistencies appeared in the part of the story in which she attempted to solve the problem; and they all involved the violation of a physical principle. Two inconsistencies were developed for each story: a falsehood and a factual contradiction. Both inconsistencies were expressed either in positive story resolutions (i.e., the problem was solved) or in negative story resolutions (i.e., the problem was not solved). Table 1 contains an example of one story in the two inconsistency types (falsehood and factual contradiction) and the two story outcomes (positive and negative).

[Insert Table 1 about here.]

In the falsehood stories, the resolution of the problem involved the negation of an unstated common sense physical law. The unstated (implicit) common sense law that was violated was, in most cases, the law of gravity. For example, in the story about the sugar and the iron filings, the sugar and the iron filings lift themselves up from the table and go back in their boxes. In another story, two labels fly in the air and out an opened window. In the remaining two stories, a key falls off a magnet to which it was attracted without any explanation, and a glass of water moves by itself to the end of a table and falls down.

In the factual contradiction stories, the resolution of the problem involved the negation of an explicitly stated, but relatively unfamiliar, scientific principle. The four principles which were explicitly contradicted were the following: (a) that magnets attract iron but not sugar, (b) that balloons filled with helium rise up but balloons filled with oxygen do not, (c) that magnets lose their magnetic qualities when heated, and (d) when a mixture of water and salt is heated, water boils and turns into steam but salt does not.

Familiarity with the principles which were violated in the text was determined in a pilot study. In this pilot study, 40 first graders (mean age 6.7), 33 third graders (mean age 8.6), and 22 fifth graders (mean age 10.5) were asked to verify the false propositions out of context. A total of 16 false propositions, corresponding to those found in the experimental materials, were used. These false propositions were divided into two lists of eight propositions each so that the same proposition did not occur more than once. Twelve additional propositions of equivalent difficulty were added to each list as foils. Thus, each final list consisted of 20 propositions, half of which were true and half were false. Within each grade, half of the children received one list and the remaining half the other. The first and third grade children read the sentences along with a teacher who read them out loud for them and then marked each sentence as true or false. The fifth grade children read and verified the sentences without any help from the teacher. The first grade children were correct 92% of the time on the falsehoods and 53% of the time on the factual contradictions. The third grade children were correct 93% on the falsehoods and 61% of the time on the factual contradictions. Finally, the fifth grade children were correct 97% and 76% of the time on the falsehoods and factual contradictions respectively. Thus, we were able to verify that the experimental materials possessed the desired characteristics, especially with the familiarity variable.

Design

The design of this study included a listening task for first graders, a reading and a listening task for third graders, and a reading task for fifth graders. Thus the listening/reading comparison was evaluated at grade 3 only, while cross-age listening comparisons were evaluated as grade 1 versus grade 3 contrasts, and cross-age reading tasks were evaluated as grade 3 versus grade 5 contrasts. There were 20 children within each grade by task condition who were separated into four blocks of five children each. The children in each block heard or read all four stories but in different combinations of inconsistency type (falsehood vs. factual contradiction) and outcome value (positive vs. negative). The order in which the children within each group heard the stories was determined by a latin square design. Thus, the full design of the study was a 3 (Grade) x 2 (Task) x 4 (Blocks) x 5 (Replications within each block) x 2 (Inconsistency Type) x 2 (Outcome Value) x 4 (Story), with inconsistency type, outcome value, and story as within-subject factors. However, analyses were only conducted across two grade levels (1-3 or 3-5) or across tasks (listening-reading) within grade 3.

Procedure

The children were tested individually. Testing took place in a private room in the school and lasted approximately 30 to 40 minutes. In the listening condition, each child was told that s/he was going to listen to four stories about a girl named Georgette who got into various kinds of trouble as she played in her father's lab. S/he was told "to listen very carefully to each story, because there is something wrong with each one of them. Something wrong with the way the author wrote them--something that doesn't make sense. We would like you to listen very carefully to each story and then tell the story back to us, and also tell us what it is that does not make any sense." In phase 1, each story was read twice by the experimenter and then the child was asked (a) to recall the story, and (b) to say what it was about the story that did not make any sense, and (c) to justify the response. The same procedure was followed for all stories. Phase 2 was conducted only for stories in which inconsistencies were not detected in Phase 1. Following the same order of presentation, the beginning of each story was summarized by the experimenter and the second half of the story was re-read once. The instructions were repeated and the child was asked once more to say what it was about the story that did not make any sense and why. In either phase 1 or 2, if more than one inconsistency was reported, the child was asked to select the most important problem and to explain why. In the reading condition, the same procedures were followed in both phases 1 and 2, except that the children read the stories on their own, rather than listening to them. All story recalls and responses to the inconsistency detection questions were tape-recorded.

Scoring

The children's recalls of the stories and responses to the inconsistency detection questions were transcribed and then scored by two independent judges. Inconsistency detection was determined on the basis of the children's verbal responses and justifications of the inconsistencies they reported. The recalls were scored by identifying story units on the basis of the work done on story comprehension (Stein & Glenn, 1978; Mandler & Johnson, 1977). These story units included a setting, an initiating event, a problem recognition, a principle, a plan, a resolution and a reaction; they are described in greater detail in Table 2. Inter-judge agreement was 94% for the first graders, 97% for the third graders, and 96% for the fifth graders. All cases of disagreement were resolved by discussion between the two judges.

[Insert Table 2 about here.]

Results and Discussion

Inconsistency Detection

Table 3 presents the frequency of inconsistency detections as a function of grade, task type, inconsistency type, and outcome value at phase 1 and at phases 1 and 2 combined.

[Insert Table 3 about here.]

These data were analyzed by performing three separate analyses of variance (one for grade 1 vs. grade 3 listening, one for grade 3 vs. grade 5 reading and one for grade 3, listening vs. reading) on the proportion of correct responses at phase 2 (actually the sum of correct inconsistency detections for phases 1 and 2).

First versus third grade (listening). The first analysis compared first and third graders in the listening task on the proportion of correct inconsistency detections. It was a 2 (Grade) x 5 (Replications within each Block) x 4 (Blocks) x 2 (Inconsistency Type: Falsehood vs. Factual Contradiction) x 2 (Outcome Value: Positive vs. Negative) x 4 (Story) analysis of variance. Grade, replication and blocks were between subject factors with the replication factor nested within grade by block. Inconsistency type, outcome value and story were within subject factors. This analysis of variance showed main effects for grade, $F(1,108) = 27.45$, $MS_e = .110$, $p < .001$, inconsistency type, $F(1,108) = 22.68, p < .001$, and outcome value, $F(1,108) = 3.63, p < .05$. The main effect for grade was due to the fact that third graders (70/80) were overall better at detecting the inconsistencies than the first graders (45/80). The main effect for inconsistency type was due to the greater number of falsehoods (64/80) than factual contradiction detections (51/80). Finally, the main effect for outcome value was the result of the greater number of inconsistency detections for stories with positive (60/80) than negative (55/80) outcomes.

Only one interaction, a grade by inconsistency type, approached a standard significance level, $F(1,108) = 3.63$, $.05 < p < .06$. The interaction effect was due to the fact that the difference in the number of correct inconsistency detections between the third graders and the first graders was greater for the contradictions (15) than for the falsehoods (10).

Third grade (listening versus reading). The second analysis of variance (using a similar mix of between and within subject factors) compared the performance of the third graders in the listening and reading tasks on the proportion of correct inconsistency detections. It was a 2 (Task: listening vs. reading) x 5 (Replications within Blocks) x 4 (Blocks) x 2 (Inconsistency Type: falsehood vs. factual contradiction) x 2 (Outcome Value: positive vs. negative) x 4 (Story) analysis of variance. The results showed no interactions; there were main effects for task $F(1,108) = 5.31$, $MS_e = .117$, $p <$

.02, and inconsistency type, $F(1,108) = 17.22, p < .001$. The main effect for task was due to the fact that there was a greater number of inconsistency detections in the listening task (70/80) than in the reading task (62/80). The main effect for inconsistency type was again the result of a greater number of falsehood (74/80) than factual contradiction detections (58/80).

Third versus fifth grade (reading). The third analysis of variance compared the performance of the third and fifth graders on the reading task again on the proportion of correct inconsistency detections. It was a 2 (Grade: third vs. fifth) x 5 (Replication within Blocks) x 4 (Blocks) x 2 (Inconsistency Type) x 2 (Outcome Value) x 4 (Story) analysis of variance. The results showed main effects for grade $F(1,108) = 6.65$, $MS_e = .113$ $p < .01$, and inconsistency type, $F(1,108) = 19.85, p < .005$, and an interaction between grade and outcome value, $F(1,108) = 4.45, p < .05$. The main effect for grade was the result of the overall better performance of the fifth graders (73/80) than the third graders (62/80). The main effect for inconsistency type was again the result of the greater number of correct responses in the falsehoods (77/80) than the factual contradictions (58/80). Finally, the interaction between grade and outcome value was the result of the greater increase with grade in the number of correct inconsistency detections for stories with positive versus negative outcomes.

Summary. To summarize the three inconsistency detection analyses, falsehoods were easier to detect than factual contradictions. Older children detected more inconsistencies than younger children, but the age difference was greater in the case of contradiction detection than falsehood detection. Finally, inconsistency detection was less frequent in the reading task than the listening task, particularly in the case of the factual contradictions. As expected, the increased processing demands of the reading task affected performance more in the case of the factual contradictions than the falsehoods.

Inconsistency Detection as a Function of Recall

Were the contradictions harder to detect because of difficulties with the representation or with the proposition comparison subcomponents of the inconsistency detection task? In order to answer this question, information from the children's recalls was analyzed. Table 4 presents the descriptive data for story units recalled as a function of grade, task and inconsistency type. The falsehoods did not have a plan and a principle, hence there are no data in these cells. There were a few instances of "inconsistency repair," i.e., instances in which the principle or the resolution were spontaneously corrected. As Table 4 shows, some children, particularly third graders, spontaneously corrected the false resolutions (i.e., instead of saying that the magnet picked up the sugar, they said that the magnet picked up the iron filings), or reversed the principle so that it would be consistent with the resolution (i.e., said that magnets are supposed to pick up sugar instead of iron filings). Predictably, such "repairs" occurred only in the case of factual contradictions.

[Insert Table 4 about here.]

The main difference in recall between inconsistency type was in the resolution category; resolutions in the factual contradictions were harder to recall than those in falsehoods. The recall of the scientific principle in the factual contradiction passages was also poor, suggesting that the children, particularly the younger ones, had difficulty representing all the information necessary for detecting the factual contradictions.

Inconsistency Detection Failures

Information from the recalls was subsequently used to establish three separate categories of inconsistency detection failures: *Repairs*, *Recall Errors* and *Comparison Errors*. In a *Repair*, the children attempted to resolve the contradiction by "correcting" either the resolution or the principle. In a *Recall* error, the information necessary for inconsistency detection (i.e., either the resolution or the principle, or both) was not recalled. If one or the other was recalled, it was categorized as a

Partial Recall error. In a *Comparison* error, all the information necessary for detecting the inconsistency was recalled. In the case of the falsehoods, a comparison error meant that the children did not compare the false propositions against existing knowledge. In the case of the factual contradictions, a comparison error meant that the children had problems comparing the two inconsistent propositions.

[Insert Table 5 about here.]

The frequency of these various error types is presented in Table 5 as a function of grade, inconsistency type and task type. In the case of the familiar falsehoods, there were no repairs and no partial recall errors (because there was no principle to be recalled). These data were subjected to various loglinear analyses after it was determined that there were no interactions with story and outcome value. A loglinear analysis tests which models fit frequency data of this sort and which are likely to be inadequate, using a Chi-square statistic. Since these models are tested against lack of fit, the larger the p value the better the model fits the data (Feinberg, 1980).

Falsehoods. A 2 (Error Type) x 2 (Phase) x 2 (Grade) loglinear analysis was performed on the falsehood data separately for the listening (grade 1 vs. grade 3) and reading (grade 3 vs. grade 5) tasks. The analysis on the listening task showed that the model that best fit the data included a main effect for grade ($X^2 = 7.61$, df = 6, $p > .30$), while the model that best fit the data on the reading task showed a main effect both for grade and phase ($X^2 = 1.12$, df = 3, $p > .70$). As can be seen in Table 5, the older children had fewer overall errors than the younger children and the frequency of their errors decreased at phase 2, as compared to phase 1.

The first grade children made both recall and comparison errors with the falsehoods. The recall errors suggested that the children had not stored all information required to evaluate the falsity of the text's resolution. The comparison errors indicated that the children had access to the relevant information but did not consider the story resolution to be "false." Seven such comparison errors occurred at Phase 1. Most of these were corrected by Phase 2. It does not appear likely that these errors involved some difficulty in comparing the false propositions against existing knowledge. A comparable group of first graders had no problems recognizing the falsity of these resolutions when presented out of context in the pilot study (92% correct). The most probable explanation is that some first grade children were reluctant to consider the story resolutions as "false" because magical solutions similar to those used in the present materials (i.e., that the sugar lifted itself from the table and went back into its box) often appear in children's stories and fables. It is possible that these children, while aware of the fact that sugar cannot literally lift itself from the table to get back into its box, considered such events possible in the story context by invoking a "magical world schema." This tendency was not present in the older children as indicated by the absence of comparison errors in that group.

Factual contradictions. Another 3 (Error Type) x 2 (Phase) x 2 (Grade) loglinear analysis was performed on the contradiction data, separately for the listening and reading tasks. The analysis of the data on the listening task showed that the model that best fit the data included a main effect for phase and an interaction between error type and grade ($X^2 = 4.63$, df = 5, $p < .30$). The main effect for phase was due to children's better performance at phase 2. The error type x grade interaction was due to the decrease in the frequency of recall errors at grade 3.

Two models fit the data on the reading task almost identically. The first included main effects for error type and phase ($X^2 = 9.09$, df = 8, $p > .30$); the second included main effects for error type and grade ($X^2 = 9.28$, df = 8, $p > .30$). The main effect for error type was due to the greater frequency of recall errors as compared to comparison errors and repairs. The phase and the grade effects were in the predictable directions; that is, older students had fewer errors than younger students and, at each grade level, students made fewer errors at Phase 2.

Across both the listening and reading analyses, most of the errors were associated with incomplete recall of the inconsistent information. The few comparison errors centered mainly around one story whose principle stated that magnets do not attract metal objects when heated. Some children questioned the truth value of this principle and said that they had never heard of magnets losing their magnetic qualities. This was one of the very few instances where the children questioned the truth of a proposition.

The Adequacy of Recall as a Measure of Story Representation

It could be argued that recall is not a good measure of story representation and that the children had encoded and stored the relevant information but did not mention it in their recalls. One way to examine this argument is to look at the correlation between inconsistency detection and recall of the resolution (and the principle in the case of the contradiction). If the children had encoded and stored the resolution but did not mention it in their story recalls, then the correlation between resolution recall and inconsistency detection should be low. Table 6 shows the frequency of inconsistency detection and recall of the relevant resolution (or the resolution and the principle in the case of the factual contradictions). As can be seen, there were only a few instances where the inconsistencies had been detected without the resolution having been recalled.Pearson correlation coefficients showed significant correlations between the two variables at all grade x condition levels.

[Insert Table 6 about here.]

Summary

Children found the unfamiliar factual contradictions harder to detect than the familiar falsehoods. Contradiction detection was also affected by differences in the children's age (the older children did better than the younger ones) and by the nature of the task (more contradictions were detected in the listening task than in the reading task). Information from the children's recall protocols suggested that the difficulties in detecting the contradictions were related more to remembering the inconsistent information (recall errors) than comparing the inconsistent propositions once they were recalled (comparison errors). Recall of the story resolution and/or the story principle was significantly correlated with inconsistency detection. It thus appears that the major source of inconsistency detection failure in this experiment was related more to children's difficulty in representing the inconsistent information in memory, rather than in comparing their representations of the inconsistent propositions to each other.

EXPERIMENT 2

Experiment 2 was designed to provide a more precise test of the source of inconsistency detection difficulty--representation versus comparison--by controlling the familiarity variable. In addition, we wanted to further test the abilities of the first grade children to evaluate texts for internal consistency. It could be argued that the first grade children had little opportunity to make comparison errors in Experiment 1 because they could not remember all the relevant information in the first place. We reasoned that young children's abilities to compare two inconsistent propositions against each other should be tested in a situation where the information contained in the two contradictory propositions presented in the text is novel enough not to be part of the children's existing knowledge, but familiar enough to be encoded without difficulty. (Recall that in Experiment 1, all falsehoods were violations of familiar principles while all factual contradictions involved violations of a relatively obscure scientific principle.)

We created materials appropriate for such a test by writing new texts that centered upon familiar events, such as cooking spaghetti. Then, for *each* text, we created an inconsistency which was expressed either as a *factual contradiction* (i.e., was explicitly negated in the text by another proposition), or as a *falsehood* (i.e., was *not* explicitly negated). Note that when the familiarity variable is controlled, the comparison difference between falsehoods and factual contradictions may disappear. A familiar factual contradiction may be evaluated either against the textual proposition that contradicts it (for internal consistency) or against existing knowledge (for truth). In order to test children's abilities to compare inconsistent propositions, a *textual contradiction* was constructed. This contradiction could be detected *only if* the children evaluated the text for internal consistency. Recall from our original description of textual contradictions that both propositions are equally likely to be true but only one can be true in a given situation. For example, in a text about Georgette cooking spaghetti, it was first asserted that Georgette "poured the spaghetti and water into a strainer." Later, this assertion was contradicted by saying that "She had not poured the spaghetti and water mixture into the strainer."

Children's detection of these *textual contradictions* were compared to their detection of a false fact (i.e., it was asserted that when spaghetti and water are poured through a strainer, the spaghetti goes through but the water does not) which was expressed either as a *factual contradiction* (i.e., it was explicitly negated earlier in the text), or as a *falsehood* (i.e., there was no prior textual proposition that negated it). All three types of inconsistencies appeared in the same story and involved the same phenomenon. Hence, any differences in inconsistency detection among these three types of inconsistencies could not be attributed to a familiarity confound.

Method

Subjects

Thirty first-grade children (ranging in age from 6.6 to 8.3, with a mean of 7.4) participated in this study. The children attended an elementary school in the same midwestern city as the children in Experiment 1. Approximately half of the children were girls and half were boys. All the children functioned at or above their grade level as determined by teacher judgment.

Materials

Two stories describing how Georgette solves a problem situation in which she gets involved were written. Both stories had the same structure: They described the creation of a problem situation and its resolution. All inconsistencies involved the violation of this resolution.

In one condition the resolution involved the negation of a true fact (*falsehood*). In a second condition the resolution involved the explicit contradiction of this same true fact (*factual contradiction*). In a third condition, the resolution contradicted an action previously described in the text as having taken place (*textual contradiction*). Table 7 presents one story in all three conditions.

Design

The children were randomly assigned to one of the three conditions: falsehood, factual contradiction, textual contradiction. There were 10 children in each condition. Each child listened to both stories. The order of presentations of each of the two stories was counterbalanced.

[Insert Table 7 about here.]

Procedure

The procedure was similar to that adopted in Experiment 1. Each subject was tested individually in the school by one of two experimenters. Testing lasted approximately 20 minutes. The experimenter read the first story twice. The instructions used were similar to those in Experiment 1. In Phase 1, the children were told to "listen carefully to the story, to tell the story back, and to say what about the story did not make sense." The same procedure was followed for the second story. This procedure was repeated a second time (Phase 2) for the stories in which inconsistencies were not detected at Phase 1.

Scoring

The children's responses to the inconsistency detection questions were scored by two independent judges following similar criteria as described for Experiment 1. Agreement was 98%; all cases of disagreement were resolved by discussion.

Results and Discussion

Inconsistency Detections

A 3 (inconsistency type) x 2 (inconsistency detection vs. nondetection) x 2 (story) loglinear analysis showed that the model that fit the data best included only a main effect for inconsistency detection ($X^2 = 1.999$, df = 10, $p > .99$), resulting from the greater frequency of detected (37) than undetected (23) inconsistencies. There were no differences in the frequency of detected inconsistencies between falsehoods (13), factual contradictions (13), and textual contradictions (11). Finally, there was no difference in the total number of inconsistencies detected between Story 1 (19) and Story 2 (18).

Inconsistency Detection Failures

Children's recall protocols were examined for error information. Three types of inconsistency detection failures were again identified: Repairs, Recall Errors and Comparison Errors.

Consistent with Experiment 1, the frequency of recall errors exceeded the other error types (11 recall errors as compared to 6 comparison errors and 6 repairs). Of the six comparison errors, half occurred in falsehoods, which did not require the comparison of two propositions.

In summary, the results of this experiment indicated that when familiarity variables is controlled, *contradictions* are not harder to detect than *falsehoods*, even by first grade children. The children detected as many *contradictions* as *falsehoods*, and there was no difference between *factual contradictions* and *textual contradictions*. These results suggest that the difficulties with factual contradiction detection in the first experiment were related more to their unfamiliarity rather than to the proposition comparison component.

Discussion

Taken together, the results of the present experiments show a relatively high rate of inconsistency detection, even for first grade children. For example, in Experiment 1, 68% of the first grade students and 93% of the third grade students detected the falsehoods. These results strengthen the findings of a few other experiments in showing that the ability to detect inconsistencies in text begins early (Ackerman, 1981; Stein & Trabasso, 1981; Wimmer, 1979). Remember, however, that we did not ask the children to simply evaluate the texts for consistency but to detect an inconsistency once informed of its presence. Pilot studies with our materials indicated that children found it much more difficult to evaluate the same texts if they were not cued to the presence of inconsistencies (see also Markman, 1979).

The hypothesis that children find it easier to evaluate a text for truth rather than for internal consistency was not supported. Falsehoods were not easier to detect than textual contradictions in Experiment 2, when the familiarity variable was controlled. It thus appears that the reason falsehoods were easier to detect than factual contradictions in Experiment 1 was because the propositions comprising the falsehoods were more familiar to the students than the contradictions, and therefore easier to represent.

The corollary hypothesis of the comparison deficit view--that children represent the inconsistent information in memory but do not detect the inconsistency for lack of adequate processing--was not supported either. There were only a few instances where the inconsistent information was recalled but the contradiction was not detected. Conversely, when these propositions were recalled, inconsistency detection was high.

While the present results are consistent with the conclusions of Wimmer (1979) and Stein and Trabasso (1981), they are at odds with the view of Flavell et al. (1981) and Markman (1977, 1979, 1985) that recall does not stand in the way of inconsistency detection. While there are too many differences between these various experiments to allow direct comparisons of the experimental findings, we will provide two possible explanations for these discrepancies.

In the Flavell, et al. (1981) experiments the measure of recall was the child's repetition of individual sentences (from a set of instructions) immediately after those sentences were read. The simple repetition of these sentences does not necessarily imply that children were representing them in long term memory (although it does show that the information was encoded in short term memory). Recall of the entire text gives a better indication of the mental representation of the passage that the subject has formed. It could, of course, be argued that it is not possible to separate the process of forming a mental representation of a text from processing requirements. Comprehension is a dynamic, constructive process which involves recursive operations such as connecting propositions and drawing inferences. To that extent, information processing limitations may well get in the way of text representation and inconsistency detection (Markman, 1981, 1985). However, it should also not be forgotten that the complexity of the constructive processing children can engage in is determined by what already exists in the knowledge base. The interaction between prior knowledge and information processing strategies is still an issue which is not well understood, and there are a number of proposals indicating that what may develop with age is not information processing ability per se but rather the amount and complexity of the knowledge base (e.g., Chi, 1978, Chi & Ceci, in press).

Markman's (1979) experiments and materials are closer to ours. In her experiments the children apparently recalled the contradictory information but failed to notice the contradictions. One possible reason for this finding is that the contradictions in Markman's texts involved counter-intuitive principles (e.g., that ice-cream does not melt when heated, that ants cannot smell, that fish cannot see the color of their food). These propositions run counter to children's existing beliefs and may have encouraged children to question their truth, rather than evaluate their consistency. Something similar happened in our text about magnets, in which the (true) fact that magnets lose their magnetic qualities when heated, was questioned by some children.

In our opinion, children's inconsistency detection failures in the case of counter-intuitive text may have less to do with their ability to compare inconsistent propositions than with their domain knowledge. It is becoming increasingly apparent that children's comprehension failures are often related to difficulty in restructuring existing knowledge which is incompatible with information presented in the text (e.g., Alverman, Smith, & Readance, 1985; Anderson & Smith, 1982; Carey, 1985; Driver & Easley, 1978; Osborne & Wittrock, 1983; Vosniadou & Brewer, 1987). When children read texts which describe information inconsistent with what they know, they tend to assimilate the information presented in the text to the already existing knowledge structures. This process of assimilation usually results in distortions and misconceptions of the text. Such text distortions are not restricted to children; they also occur in adults (diSessa, 1982; White, 1983; McCloskey, 1983).

To conclude, the results of the present experiments indicated that children's inconsistency detection failures were more related to difficulty in representing the inconsistent information in memory than comparing the representations of the inconsistent proportions to each other. Thus, the hypothesis that children have particular difficulty in comparing inconsistent propositions to each other was not supported in these experiments. These results suggest that greater attention should be paid to how children's mental representations of a text affect inconsistency detection and comprehension monitoring. Children often lack the background knowledge necessary to form an adequate mental representation. Or, they may have beliefs or theories which are contrary to the information included in the text. Inadequate and/or incompatible prior knowledge may result in distortions of the text information, failures to detect inconsistencies, repair of inconsistencies, or a disposition toward evaluating the text for truth rather than for consistency. Future research should focus on the interaction between children's prior knowledge and their inconsistency detection failures.

References

Ackerman, B. P. (1981). The understanding of young children and adults of the deictic adequacy of communications. *Journal of Experimental Child Psychology, 31*, 256-270.

Alverman, D. E., Smith, L. C., Readence, J. E. (1985). Prior knowledge activation and the comprehension of compatible and incompatible text. *Reading Research Quarterly, 20*(4), 420-436.

Anderson, C. W., & Smith, E. L. (1982, April). *Student conceptions of light, color and seeing: Phase 1 test results*. Paper presented at the annual convention of the National Association for Research on Science Teaching, Fontana, WI.

August, D. L., Flavell, J. H., & Clift, R. (1984). Comparison of comprehension monitoring of skilled and less skilled readers. *Reading Research Quarterly, 20*(1), 39-45.

Baker, L. (1979). Comprehension monitoring: Identifying and coping with text confusions. *Journal of Reading Behavior, 11*, 366-374.

Baker, L., & Anderson, R. I. (1982). Effects of inconsistent information on text processing: Evidence for comprehension monitoring. *Reading Research Quarterly, 17*, 281-294.

Bransford, J. D., & Johnson, M. K. (1973). Consideration of some problems in comprehension. In W. Chase (Ed.), *Visual information processing*. New York: Academic Press.

Brown, A. L., Smiley, S. S., Day, T. D., Townsend, M. A. R., & Lawton, S. C. (1977). Intrusion of a thematic idea in children's comprehension and retention of stories. *Child Development, 48*, 1454-1466.

Carey, S., (1985). *Conceptual change in childhood*. Cambridge, MA: MIT Press.

Chi, M. T. H. (1978). Knowledge structures and memory development. In R. S. Siegler (Ed.), *Children's thinking: What develops*? Hillsdale, NJ: Erlbaum.

Chi, M. T. H., & Ceci, S. J. (in press). In H. W. Reese (Ed.), *Advances in child development and behavior* (Vol. 20).

diSessa, A. (1982). Unlearning Aristotelian physics: A study of knowledge based learning. *Cognitive Science, 6*, 37-75.

Driver, R., & Easley, J. (1978). Pupils and paradigms: A review of literature related to concept development in adolescent science students. *Studies in Science Education, 5*, 61-84.

Elliot-Faust, D. J., & Pressley, M. (1986). How to teach comparison processing to increase children's short- and long-term listening comprehension monitoring. *Journal of Educational Psychology, 78*(1), 27-33.

Feinberg, S. E. (1980). *The analysis of cross-classified categorical data*. Cambridge, MA: MIT Press.

Flavell, J. H. (1981). Cognitive monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills*. New York: Academic Press.

Flavell, J. H., Speer, J. R., Green, F. L., & August, D. (1981). The development of comprehension monitoring and knowledge about communication. *Monographs of the Society for Research in Child Development.*

Garner, R. (1980). Monitoring of understanding: An investigation of good and poor readers' awareness of induced miscomprehension of text. *Journal of Reading Behavior, 12*, 55-64.

Garner, R. (in press). Resolving comprehension failure through text lookbacks: Direct training and practice effects among good and poor comprehenders in grades six and seven. *Reading Psychology.*

Gelman, R. (1978). Cognitive development. *Annual Review of Psychology, 29*, 297-332.

Ironsmith, M., & Whitehurst, G. J. (1978). The development of listener abilities in communications: How children deal with ambiguous information. *Child Development, 49*, 348-352.

Kotsonis, M. E., & Patterson, C. F. (1980). Comprehension-monitoring skills in learning disabled children. *Developmental Psychology, 16*, 541-542.

Mandler, J., & Johnson, N. (1977). Remembrance of things parsed: Story structure and recall. *Cognitive Psychology.*

Markman, E. M. (1977). Realizing that you don't understand: A preliminary investigation. *Child Development, 43*, 986-992.

Markman, E. M. (1979). Realizing that you don't understand: Elementary school children's awareness of inconsistencies. *Child Development, 50*, 643-655.

Markman, E. M. (1981). Comprehension monitoring. In W. P. Dickson (Ed.), *Children's oral communication skills.* New York: Academic Press.

Markman, E. M. (1985). Comprehension monitoring: Developmental and educational issues. In S. F. Chipman, J. W. Segal, & R. Glaser (Eds.), *Thinking and learning skills, Vol. 2: Research and open questions* (pp. 275-291).

Markman, E. M., & Gorin, L. (1981). Children's ability to adjust their standards for evaluating comprehension. *Journal of Educational Psychology, 73*, 320-325.

McCloskey, M. (1983). Naive theories of motion. In D. Gentner & A. L. Stevens (Eds.), *Mental models* (pp. 299-324). Hillsdale, NJ: Erlbaum.

Osborne, R. J., & Wittrock, M. C. (1983). Learning science: A generative process. *Science Education, 67*, 489-508.

Patterson, C. J., Cosgrove, J. M., & O'Brien, R. G. (1980). Nonverbal indicants of comprehension and noncomprehension in children. *Developmental Psychology, 16*, 38-48.

Pearson, P. D., & Fielding, L. (1983). Instructional implications of listening comprehension research. *Language Arts, 59*(6), 617-629.

Raphael, T. E., & Tierney, R. J. (1980). The influence of topic familiarity and the author-reader relationship on detection of inconsistent information. In M. Kamil (Ed.), *Directions in reading: Research and instruction* (Thirtieth Yearbook of the National Reading Conference). San Diego, CA: NRC.

Robinson, E. J. (1981). The child's understanding of inadequate messages and communication failure: A problem of ignorance or egocentrism? In W. P. Dickson (Ed.), *Children's oral communication skills*. New York: Academic Press.

Robinson, E. J., & Robinson, W. P. (1976a). The young child's understanding of communication. *Developmental Psychology, 12*, 328-333.

Robinson, E. J., & Robinson, W. P. (1976b). Developmental changes in the child's explanation of communication failure. *Australian Journal of Psychology, 28*, 155-165.

Robinson, E. J., & Robinson, W. P. (1977). Children's explanation of communication failure and the inadequacy of the misunderstood message. *Developmental Psychology, 13*, 156-161.

Stein, N. L., & Glenn, C. G. (1978). An analysis of story comprehension in elementary school children. In R. Freedle (Ed.), *Multidisciplinary approaches to discourse comprehension*. Hillsdale, NJ: Erlbaum.

Stein, N. L., & Trabasso, T. (1981). *What's in a story: An approach to comprehension and instruction* (Tech. Rep. No. 100). Urbana: University of Illinois, Center for the Study of Reading.

Sticht, T. G., Beek, L. J., Hanke, R. N., Kleiman, G. M., & Mamee, J. H. (1974). *Auding and reading: A developmental model*. Alexandria, VA: Human Resources Research Organization.

Vosniadou, S., & Brewer, W. F. (1987). Theories of knowledge restructuring in development. *Review of Educational Research, 57*(1), 51-67.

White, B. Y. (1983). Sources of difficulty in understanding Newtonian dynamics. *Cognitive Science, 7*, 41-65.

Whitehurst, G. J. (1981). Commentary. *Monographs of the Society for Research in Child Development*.

Winograd, P., & Johnston, P. (1980). *Comprehension monitoring and the error detection paradigm* (Tech. Rep. No. 153). Urbana: University of Illinois, Center for the Study of Reading.

Wimmer, H. (1979). Processing of script deviations by young children. *Discourse Processes, 2*, 301-310.

Table 1

Example of One Story Used in the Different Inconsistency Type and Outcome Value Conditions

Today Georgette came from school and found a note for her on the kitchen table.

Georgette,

We are out of sugar, and I need some for an experiment. Would you borrow a cup of sugar from Mr. Johnson, next door, and put it in the dish I have left on my desk. I'll be home about 5:30.

Thanks,

Dad

Georgette went quickly to Mr. Johnson's house and got the sugar. She put the sugar in the dish on one of her Dad's big work tables. Then she climbed upon the table to get a better look at the array of bottles and boxes on the shelves above the table. She hadn't even started exploring when a box of black stuff toppled off the shelf right onto her dish of sugar. Georgette was very worried. Without thinking too carefully, she ran to the kitchen and got a teaspoon. She dipped the spoon into the dish, trying carefully to pick up only the black stuff. But all she managed to do was mix up the black stuff with the sugar even more.

"This'll never work," she thought as she watched the mixture look more and more like salt and pepper all mixed up.

When she was almost ready to give up, she noticed a label on the box she had spilled. It said: "IRON FILINGS." She wasn't sure what filings were, but she did recognize the word "iron."

<u>Falsehood and Positive Outcome Value</u>

"I wish something could happen to solve this problem." She stayed there for a while looking at the sugar and iron filing mess on the table when she saw the iron filings and the sugar slowly lifting themselves from the table and separating. The sugar went back into the measuring cup and the iron filings went back into their box. Georgette was very relieved that everything was back in place.

<u>Falsehood and Negative Outcome Value</u>

"I wish something could happen to solve this problem," she said to herself. She stayed there for a while looking at the sugar and the iron filing mess on the table when she saw the iron filings and the sugar slowly lifting themselves from the table and falling all over the floor. There was no way Georgette could separate them now. Georgette was very worried about what happened and didn't know what she was going to tell her father.

Table 1 (Continued)

Factual Contradiction and Positive Outcome Value

An idea came to her.

"If it's iron," she thought, "I know how to get it out. All I need is a magnet. Magnets pick up iron but not sugar."

She ran upstairs to her room and got a magnet. She dipped the magnet into the funny looking salt-and-pepper-like mess in the dish. The magnet picked up the sugar, leaving the iron filings on the table.

Georgette was very relieved and happy that everything was back in place. She put the iron filings in their box and placed the box on the shelf exactly where she had found it.

Factual Contradiction and Negative Outcome Value

"If it's iron," she thought, "I know how to get it out. All I need is a magnet. Magnets pick up iron but not sugar."

She ran upstairs to her room and got a magnet. She dipped the magnet into the funny looking salt-and-pepper-like mess in the dish. The magnet picked up the iron filings and the sugar, leaving nothing on the table.

Georgette was very unhappy that the problem was not solved. She had to wait for her father to come and she was sure he would not like it.

Table 2

Story Units Used to Score the Iron Filings Passage

A: SETTING

Georgette came home, found note, . . . etc.

B: EPISODE

1. Initiating Event

A box of black stuff fell into the bowl of sugar.

2. Problem Recognition

Anything that indicates that she knew there was a problem

Falsehood		Factual Contradiction	
3. Plan		3. Plan	
			Need magnet
4. Principle		4. Principle	
			Magnets pick up iron but not sugar.
5. Resolution		5. Resolution	
(positive)	The sugar and iron filings lifted themselves off the table and separated.	(positive)	Magnet picked up the sugar leaving the iron filings on the table.
(negative)	The sugar and iron filings lifted themselves off the table, fell on the floor and there was no way to separate them.	(negative)	The magnet picked up the iron filings and the sugar leaving nothing on the table.
6. Reaction		6. Reaction	
(positive)	Georgette was happy.	(positive)	Georgette was happy.
(negative)	Georgette was worried.	(negative)	Georgette was worried.

Table 3

Frequency of Inconsistency Detections as a Function of Grade, Task Type, Inconsistency Type and Outcome Value, at Phases 1, and 1 and 2 Combined (out of 20)

	Falsehood					Contradiction					
	Phase 1		Phase 1 & 2		Sub-Total	Phase 1		Phase 1 & 2		Sub-Total	Total
	Pos.	Neg.	Pos.	Neg.		Pos.	Neg.	Pos.	Neg.		
Listening											
Grade 1	12	11	13	14	27/40	6	8	10	8	18/40	45/80
Grade 3	18	17	19	18	37/40	12	7	18	15	33/40	70/80
TOTAL					64/80					51/80	115/160
Reading											
Grade 3	14	15	17	20	37/40	5	6	12	13	25/40	62/80
Grade 5	19	18	20	20	40/40	15	10	19	14	33/40	73/80
TOTAL					77/80					58/80	135/80

Table 4

Frequency of Story Units Recalled as a Function of Grade, Task Type and Inconsistency Type (out of 40)

Task	Initiating Event		Problem Recognition		Plan		Principle		Resolution		Reaction		Principle Reversed		Resolution Reversed	
	False	Contr.	False	Contr.	False	Contr.	False	Contr.	False	Contr.	False	Contr.	False	Contr.	False	Contr.
Listening																
Grade 1	35	31	23	22	–	31	–	8	25	20	7	15	–	0	0	2
Grade 3	39	40	37	31	–	39	–	24	35	28	19	24	–	1	0	10
Reading																
Grade 3	37	37	31	31	–	37	–	16	29	25	13	15	–	3	0	5
Grade 5	39	39	32	37	–	36	–	29	36	30	4	16	–	1	0	5

Table 5

Frequency of Error Types as a Function of Grade, Task Type and Inconsistency Type

	Phase 1					Phase 2				
	Repair	Full Recall	Partial Recall	Comparison	Total	Repair	Full Recall	Partial Recall	Comparison	Total
Falsehoods										
Listening										
Grade 1	0	10	--	7	17	0	9	--	4	13
Grade 3	0	3	--	2	5	0	2	--	1	3
Reading										
Grade 3	0	10	--	1	11	0	3	--	0	3
Grade 5	0	1	--	2	3	0	0	--	0	0
Factual Contradictions										
Listening										
Grade 1	2	13	8	3	26	2	10	7	3	22
Grade 3	8	0	8	5	21	3	0	1	3	7
Reading										
Grade 3	7	6	11	5	29	3	4	5	3	15
Grade 5	4	0	8	3	15	0	0	5	2	7

Table 6

<u>Falsehoods</u>

Frequency of

Resolution Recall/Inconsistency Detection (out of 40)

Listening Task	
Grade 1	25/27
Grade 3	35/37
Reading Task	
Grade 3	29/37
Grade 5	36/40

<u>Factual Contradictions</u>

Frequency of

Resolution and/or Principle Recall/Inconsistency Detection (out of 40)

Listening Task	
Grade 1	16/18
Grade 3	33/33
Reading Task	
Grade 3	25/25
Grade 5	30/33

Table 7

Example of One Story Used in the Three Inconsistency Conditions

On Tuesday evening, Georgette came home from her friend's house around 6:00. Pretty soon she started to get hungry. Her father was still down in the lab working, so Georgette decided to surprise him and make dinner herself. She thought she would try to make spaghetti because she had watched her father make that before and it didn't look too hard.

First Georgette filled a pot with water and put it on the stove. She turned on the flame and soon the water was boiling, so she put a boxful of spaghetti into the water. The water boiled again and Georgette watched it until the spaghetti got soft and looked like it was done. Now she had to think of a way to take the spaghetti out of the water.

<u>Falsehood</u>

Then she remembered that her father had used a strainer to separate spaghetti from water. Georgette looked in the kitchen cabinets and found a strainer. She put the strainer over a bowl. She then poured the spaghetti and water into the strainer. As she did this, the spaghetti passed through the holes of the strainer into the bowl and the water stayed in the strainer. Georgette was glad that the water and spaghetti were separated. She set the bowl of spaghetti on the counter. Then she warmed up some spaghetti sauce, mixed the sauce and spaghetti together, and called her father to dinner.

<u>Factual Contradiction</u>

Then she remembered that her father had used a strainer to separate spaghetti from water. When you pour spaghetti and water into a strainer, the water passes through the holes and the spaghetti stays in the strainer.

So Georgette looked in the kitchen cabinets and found a strainer. She put the strainer over a bowl. She then poured the spaghetti and water into the strainer. As she did this, the spaghetti passed through the holes of the strainer into the bowl and the water stayed in the strainer. Georgette was glad that the water and spaghetti were separated. She set the bowl of spaghetti on the counter. Then she warmed up some spaghetti sauce, mixed the sauce and spaghetti together, and called her father to dinner.

<u>Textual Contradiction</u>

Then she remembered that her father had used a strainer to separate spaghetti from water. So Georgette looked in the kitchen cabinets and found a strainer. She put the strainer over a bowl, and poured the spaghetti and water into the strainer. The water went into the bowl leaving the strainer full of spaghetti.

Georgette was glad the spaghetti did not have any water in it anymore. She put the spaghetti back into the pot, and looked for some spaghetti sauce. She found a bottle of spaghetti sauce and warmed it up. When the sauce was ready, she noticed that she had not poured the spaghetti and water mixture into the strainer yet. So, she poured it into the strainer. And when all the water was gone, she mixed the sauce and the spaghetti together, and called her father to dinner.

This page is intentionally blank.